

# Comportamiento de parámetros del patrón respiratorio en clasificadores para la predicción del proceso *weaning*

J. Chaparro Preciado<sup>1</sup>, B.F. Giraldo Giraldo<sup>2,3,4</sup>, P. Caminal<sup>3,4</sup>, S. Benito<sup>5</sup>

<sup>1</sup>Escuela Colombiana de Ingeniería, Grupo de investigación Bioeci, Bogotá, Colombia, {Javier.chaparro}@escuelaing.edu.co

<sup>2</sup>Institut de Bioenginyeria de Catalunya (IBEC), Barcelona, España, {bgiraldo}@ibecbarcelona.eu

<sup>3</sup>Depto. ESAIL, Universitat Politècnica de Catalunya (UPC), Barcelona, España

<sup>4</sup>CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

<sup>5</sup>Hospital de la Santa Creu i Sant Pau. Barcelona, España

## Resumen

*El proceso de destete de pacientes en unidades de cuidados intensivos es un procedimiento clínico complejo. 153 pacientes en proceso de extubación fueron estudiados: 94 pacientes con éxito en la extubación (grupo S), 38 pacientes que no lograron mantener la respiración espontánea y fueron reconectados al ventilador (grupo F), y 21 pacientes con éxito en la prueba, pero antes de 48 horas tuvieron que ser reintubados (grupo R). El patrón respiratorio de cada paciente se caracterizó a partir de las series temporales: tiempo de inspiración ( $T_I$ ), tiempo de espiración ( $T_E$ ), duración del ciclo respiratorio ( $T_{Tot}$ ), volumen corriente ( $V_T$ ), fracción inspiratoria ( $T_I / T_{Tot}$ ), flujo medio inspirado ( $V_T / T_I$ ), y el índice de ventilación superficial ( $f / V_T$ ), siendo  $f$  la frecuencia respiratoria. Los parámetros más relevantes del patrón respiratorio se obtuvieron usando modelos autorregresivos (AR), modelos autoregresivos de media móvil (ARMA) y modelos autorregresivos con entrada exógena (ARX). Se ha propuesto la evaluación de estos parámetros utilizando clasificadores como regresión logística (LR), análisis discriminante lineal (LDA), máquinas de soporte vectorial (SVM) y clasificación mediante un árbol de decisión (CART) para discriminar entre los pacientes de los grupos S, F y R. Los mejores resultados se obtuvieron con el clasificador CART, con una exactitud del 93% (98% de sensibilidad y 82% especificidad).*

## 1. Introducción

Uno de los problemas más difíciles en las unidades de cuidados intensivos es el proceso de desconexión de la ventilación mecánica, denominada destete. A pesar de los avances en la ventilación mecánica y el soporte respiratorio, la determinación del momento óptimo de extubación es aún muy imprecisa [1]. Cuando la ventilación mecánica es desconectada hasta un 25% de pacientes tienen dificultad para mantener la respiración espontánea necesitando ser reconectados al ventilador.

El número de pacientes que tienen que ser reintubados antes de 48 h, previamente extubados por considerar exitosa la prueba de destete representa un porcentaje superior al 10% [2]. Esto hace necesario una mayor precisión en el momento óptimo de extubación y todo lo relacionado con el destete, una vez se han superado las causas clínicas que han llevado a la intubación [3], [4].

Varios estudios reportan que aproximadamente el 40% de los pacientes de las unidades de cuidados intensivos

necesitan ventilación mecánica. Entre ellos, el 90% pueden ser desconectados del ventilador en pocos días. Los pacientes reintubados requieren mayor tiempo de hospitalización y tienen una mayor tasa de mortalidad [5].

Muchos criterios para destete basados en índices como el volumen minuto, la máxima presión inspiratoria, el volumen corriente, el índice de relación superficial, y parámetros como compliancia, resistencia, oxigenación y presión, han sido definidos para determinar si un paciente es óptimo para la extubación. La mayoría de estas pruebas son sensibles, pero no son específicas, por lo que aunque un paciente no cumpla con los criterios propios del destete, éste podría ser desconectado. Se ha demostrado que los índices evaluados son pobres predictores del momento del destete en pacientes en unidad de cuidados intensivos [6], [7].

El patrón respiratorio que describe la función mecánica del sistema pulmonar se puede caracterizar a partir de las series temporales: tiempo de inspiración ( $T_I$ ), tiempo de espiración ( $T_E$ ), duración del ciclo respiratorio ( $T_{Tot}$ ), volumen circulante ( $V_T$ ), fracción inspiratoria ( $T_I / T_{Tot}$ ), flujo inspirado medio ( $V_T / T_I$ ) y el índice de ventilación superficial ( $f / V_T$ ), siendo  $f$  la frecuencia respiratoria. La variabilidad del patrón respiratorio puede ser estudiada mediante modelos autorregresivos (AR), modelos autorregresivos de media móvil (ARMA), y modelos autorregresivos con entrada exógena (ARX). Los parámetros más relevantes se obtuvieron en [8]. Este trabajo propone usar clasificadores como la regresión logística (LR), el análisis discriminante lineal (LDA), las máquinas de soporte vectorial (SVM), y la clasificación mediante un árbol de decisión (CART), para obtener aquellos parámetros que permitan diferenciar entre los grupos de pacientes éxito (grupo S), fracaso (grupo F), y reintubados (grupo R).

## 2. Metodología

### 2.1. Pacientes

Se registraron señales de flujo respiratorio de 153 pacientes (base de datos WEANDB). Los pacientes fueron registrados en las unidades de cuidados intensivos de los hospitales Santa Creu i Sant Pau de Barcelona, y

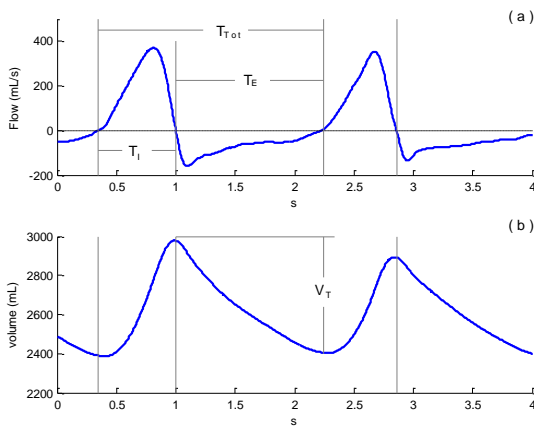
Getafe de Getafe, de acuerdo con los protocolos aprobados por los comités de éticos locales.

Los pacientes fueron sometidos a prueba de tubo en T, desconectando del ventilador del tubo endotraqueal y manteniendo la respiración espontánea durante 30 min. Si los pacientes mantuvieron la respiración espontánea durante la prueba fueron extubados, en caso contrario fueron reconectados. Cuando los pacientes pasadas 48h mantenían aún la respiración espontánea, el proceso de destete era considerado exitoso, si no, los pacientes fueron reintubados.

Los pacientes fueron clasificados en tres grupos: grupo S, 94 pacientes (61 varones, 33 mujeres, edad  $65 \pm 17$  años) con éxito en el destete; grupo F, 38 pacientes (24 varones, 14 mujeres, edad  $67 \pm 15$  años) que no pudieron mantener la respiración espontánea, y grupo R, 21 pacientes (11 varones, 10 mujeres, edad  $68 \pm 14$  años) que tuvieron éxito en la prueba de destete, pero necesitaron ser reintubados antes de 48 h.

## 2.2. Caracterización del patrón respiratorio

La señal de flujo respiratorio se obtuvo mediante un pneumotacógrafo (monitor Datex-Ohmeda transductor de reluctancia variable) conectado a un tubo endotraqueal. Las señales fueron registradas a 250 Hz, durante 30 min. La señal respiratoria se procesó para obtener las siguientes series temporales tiempo siguiente, para cada paciente:  $T_I$ ,  $T_E$ ,  $T_{Tot}$ ,  $V_T$ ,  $T_I / T_{Tot}$ ,  $V_T / T_I$ , frecuencia respiratoria ( $f$ ) y  $f / VT$  (Figura 1).



**Figura 1.** (a) señal de flujo respiratorio y sus series temporales: tiempo de inspiración ( $T_I$ ), tiempo de espiración ( $T_E$ ) y duración del ciclo respiratorio ( $T_{Tot}$ ). (b) el volumen de la señal respiratoria y volumen corriente ( $V_T$ ).

## 2.3. Técnicas de modelado

Aplicando modelos como AR, ARMA y ARX para cada serie temporal, características como el orden del modelo, el primer coeficiente y el error de predicción final (FPE) se obtuvieron para cada paciente [8]. Se seleccionaron los parámetros más relevantes que permita diferenciar entre los grupos de pacientes. Se aplicaron las pruebas estadísticas de Kruskal-Wallis y Mann-Whitney al comparar tres y dos grupos de pacientes, respectivamente.

## 2.4. Métodos de clasificación

Se aplicaron los siguientes métodos para la clasificación de los tres grupos de pacientes. Se utilizó la validación cruzada para validar estos clasificadores [9].

– *Regresión logística.* La regresión logística (RL) es un tipo de análisis de regresión utilizado para predecir el resultado de una variable categórica basado en una o más variables predictoras. El modelo de regresión logística de  $k$ -variables puede definirse como [10]

$$p = \frac{1}{1 + e^{-(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k)}} \quad (1)$$

siendo  $p$  la probabilidad de ocurrencia de un evento  $x$  de una serie de datos  $X$ , y  $\alpha_k$  los pesos de los parámetros.

– *Análisis discriminante lineal.* El análisis discriminante lineal (LDA) se utiliza para encontrar una combinación lineal de parámetros que clasifiquen dos o más categorías. Este método maximiza el ratio de las varianzas entre grupos y dentro de cada grupo, garantizando la máxima separación [11]. Se puede definir como

$$Y = \mu_0 + \mu_1 X_1 + \mu_2 X_2 + \dots + \mu_k X_k \quad (2)$$

donde  $X_i$  se corresponde con los parámetros independientes,  $\mu_0$  con el término independiente, y  $\mu_i$  es el coeficiente de la función discriminante

– *Máquinas de soporte vectorial.* Las máquinas de soporte vectorial (SVM) se basa en la transformación de los datos a un espacio dimensional superior para convertir un problema de clasificación complejo en uno más simple que se puede resolver mediante una función discriminante lineal, conocida como hiperplano, definido por [12]

$$f(x) = wz + b = \sum_i^L \alpha_i y_i K(x_i, x_j) + b \quad (3)$$

donde  $a_i$  y  $b$  se determinan para resolver el problema de mayor escala, para obtener los valores óptimos globales [13], [14].

– *Árbol de decisión.* Un árbol de decisión es un método de uso común en la de minería de datos y se utiliza como un modelo predictivo. El método de árbol de clasificación y regresión (CART) consiste en examinar todas las posibles combinaciones de todas las variables incluidas en el análisis. Los resultados se presentan en forma de árbol invertido, comenzando con un nodo raíz del que se generan nodos descendientes a partir de reglas de decisión, formando un diagrama secuencial que muestra los posibles resultados. Los nodos terminales presentan los valores alcanzados para una decisión final [15], [16].

## 2.5. Evaluación de las medidas

La calidad de los resultados son analizados en función de la exactitud ( $Acc$ ), la sensibilidad ( $Sn$ ) y la especificidad ( $Sp$ ) para cada caso. Estas medidas se construyen a partir de una matriz de confusión, que muestra una clasificación binaria donde  $tp$  (los valores) son verdaderos positivos,  $fp$  falsos positivos,  $tn$  verdaderos negativos, y  $fn$  falsos negativos [17]. Estas medidas están definidas por

$$Acc = \frac{t_p + t_n}{t_p + f_p + t_n + f_n} ; Sn = \frac{t_p}{t_p + f_n} ; Sp = \frac{t_n}{f_n + t_n} \quad (4)$$

### 3. Resultados

Los parámetros que presentaron diferencias estadísticamente significativa, al comparar los 3 grupos de pacientes (S, F y R), fueron el tiempo de espiración ( $T_E$ ), el tiempo de inspiración ( $T_I$ ), la duración del ciclo respiratorio ( $T_{Tot}$ ) y el índice de ventilación superficial ( $f/V_T$ ). Aplicando los modelos autorregresivos (AR), los mejores resultados se obtuvieron con el orden del modelo y el primer coeficiente.

Con los modelos ARMA (p, q), los valores medios del orden q de las series temporales  $T_{Tot}$ ,  $T_I$  y  $T_E$  presentaron una tendencia mayor en el grupo S que en el grupo F, siendo  $T_E$  el parámetro más relevante ( $p = 0.02$ ). Los resultados más relevantes se obtuvieron aplicando el modelo ARMA a las series temporales  $T_{Tot}$ ,  $T_I$ ,  $T_E$  y  $f/V_T$ . Se obtuvieron diferencias significativas con el error de predicción final (FPE) entre los tres grupos de pacientes.

En resumen, el tiempo de espiración, el tiempo de inspiración, la duración del ciclo respiratorio y el índice de ventilación superficial son las series temporales que presentaron las mejores diferencias entre los grupos de pacientes S, F y R a partir del valor medio, del orden modelo AR, del primer coeficiente del modelo AR, y de FPE del modelo ARMA (Tabla 1). La calidad de estas series para diferenciar los grupos de pacientes se evaluó con los clasificadores propuestos en este trabajo.

| Serías temporales                             | Parámetros                   | ID       | p-valor  |
|---|------------------------------|----------|----------|
| Tiempo de espiración ( $T_E$ )                | Media                        | $X_1$    | < 0.0001 |
|   | Orden modelo AR              | $X_2$    | < 0.001  |
|   | Primer coeficiente modelo AR | $X_3$    | < 0.005  |
|   | FPE modelo ARMA              | $X_4$    | < 0.01   |
| Tiempo de inspiración ( $T_I$ )               | Media                        | $X_5$    | < 0.0005 |
|   | Orden modelo AR              | $X_6$    | < 0.001  |
|   | Primer coeficiente modelo AR | $X_7$    | < 0.005  |
|   | FPE modelo ARMA              | $X_8$    | < 0.005  |
| Duración del ciclo respiratorio ( $T_{Tot}$ ) | Media                        | $X_9$    | < 0.0005 |
|   | Orden modelo AR              | $X_{10}$ | < 0.001  |
|   | Primer coeficiente modelo AR | $X_{11}$ | < 0.005  |
|   | FPE modelo ARMA              | $X_{12}$ | < 0.01   |
| Índice de ventilación superficial ( $f/V_T$ ) | Media                        | $X_{13}$ | < 0.05   |
|   | Orden modelo AR              | $X_{14}$ | < 0.01   |
|   | Primer coeficiente modelo AR | $X_{15}$ | < 0.01   |
|   | FPE modelo ARMA              | $X_{16}$ | < 0.05   |

**Tabla 1.** Los parámetros más relevantes que caracterizan el patrón respiratorio, obtenidos con modelos autorregresivos.

Los mejores resultados se obtuvieron utilizando todos los parámetros relacionados con  $T_E$  (media, primer coeficiente del modelo AR, FPE del modelo ARMA),

para los cuatro clasificadores, con una exactitud alrededor del 86%, una sensibilidad superior al 86% y una especificidad superior al 70% en la mayoría de los casos (Tabla 2).

Los resultados de los clasificadores con  $T_I$  presentaron valores de exactitud, sensibilidad y especificidad inferiores a los obtenidos con  $T_E$ .

La Tabla 3 presenta los resultados obtenidos con  $T_{Tot}$ , siendo mejores que los obtenidos con  $T_I$ , pero inferiores a los de  $T_E$ . La mejor exactitud, sensibilidad y especificidad se obtuvo con el método del árbol de decisión.

Las series temporales de  $f/V_T$  mostraron valores cercanos al 90% para la exactitud y la sensibilidad, pero con baja especificidad en todos los clasificadores.

| Método               | Acc  | Sn   | Sp   |
|----------------------|------|------|------|
| Regresión logística  | 0.86 | 0.92 | 0.70 |
| Discriminante lineal | 0.86 | 0.90 | 0.74 |
| SVM                  | 0.85 | 0.95 | 0.61 |
| Árbol decisión       | 0.85 | 0.87 | 0.79 |

**Tabla 2.** Exactitud, sensibilidad y especificidad obtenidas con los 4 clasificadores para el tiempo de espiración  $T_E$ .

| Método               | Acc  | Sn   | Sp   |
|----------------------|------|------|------|
| Regresión logística  | 0.71 | 0.85 | 0.37 |
| Discriminante lineal | 0.72 | 0.77 | 0.53 |
| SVM                  | 0.72 | 0.88 | 0.34 |
| Árbol decisión       | 0.83 | 0.93 | 0.58 |

**Tabla 3.** Exactitud, sensibilidad y especificidad obtenidas con los 4 clasificadores para la duración del ciclo respiratorio  $T_{Tot}$ .

Finalmente se evaluaron todas las posibles combinaciones de los parámetros obtenidos en la Tabla 1, para todos los clasificadores. Los mejores resultados se obtuvieron con la combinación de la media de  $T_E$  ( $X_1$ ), el primer coeficiente del modelo AR de  $T_E$  ( $X_3$ ) y el primer coeficiente del modelo AR de  $f/V_T$  ( $X_{15}$ ) (Tabla 4).

| Método               | Acc  | Sn   | Sp   |
|----------------------|------|------|------|
| Regresión logística  | 0.89 | 0.94 | 0.75 |
| Discriminante lineal | 0.88 | 0.93 | 0.80 |
| SVM                  | 0.88 | 0.96 | 0.66 |

**Tabla 4.** Exactitud, sensibilidad y especificidad obtenidas con 3 clasificadores considerando la media de  $T_E$  ( $X_1$ ), el primer coeficiente del modelo AR de  $T_E$  ( $X_3$ ), y el orden del modelo AR de  $f/V_T$  ( $X_{14}$ ).

Los mejores resultados se obtuvieron con el árbol de decisión, considerando la media de  $T_E$  ( $X_1$ ), el orden del modelo AR de  $T_E$  ( $X_2$ ), de  $T_{Tot}$  ( $X_{10}$ ), y de  $f/V_T$  ( $X_{14}$ ), con una exactitud de 93%, 98% de sensibilidad y 82% de especificidad. La Tabla 5 muestra las características más relevantes y sus valores con éste método de clasificación.

| Nodo | S (%) | F (%) | Total (%) | Promotico | Nodo primario | Param.   | $p^1$ | Valores segmento |
|------|-------|-------|-----------|-----------|---------------|----------|-------|------------------|
| 0    | 71.2  | 28.8  | 100       | S         | 0             | $X_2$    | 0.000 |                  |
| 1    | 5.3   | 94.7  | 14.4      | F         | 0             | $X_2$    | 0.000 | $\leq 1.0$       |
| 2    | 82.3  | 17.7  | 85.6      | S         | 2             | $X_{14}$ | 0.000 | $> 1.0$          |
| 3    | 100   | 0.0   | 39.4      | S         | 2             | $X_{14}$ | 0.000 | $\leq 10.0$      |
| 4    | 42.9  | 57.1  | 10.6      | F         | 2             | $X_{14}$ | 0.000 | (10.0 14.0]      |
| 5    | 74.5  | 25.5  | 35.6      | E         | 2             | $X_1$    | 0.000 | $> 14.0$         |
| 6    | 0.0   | 100   | 5.3       | F         | 4             | $X_1$    | 0.018 | $\leq 1.136$     |
| 7    | 85.7  | 14.3  | 5.3       | S         | 4             | $X_1$    | 0.018 | $> 1.136$        |
| 8    | 92.3  | 7.7   | 19.7      | S         | 5             | $X_{10}$ | 0.004 | $\leq 12.0$      |
| 9    | 14.3  | 85.7  | 5.63      | F         | 5             | $X_{10}$ | 0.004 | (12.0 31.0]      |
| 10   | 71.4  | 28.6  | 10.6      | S         | 5             | $X_{10}$ | 0.004 | $> 31.0$         |

<sup>1</sup>corrección Bonferroni

**Tabla 5.** Características del árbol de decisión usando la media de  $T_E (X_1)$ , el orden del modelo AR de  $T_E (X_3)$ ,  $T_{Tot} (X_{10})$  y  $f/V_T (X_{14})$ .

#### 4. Conclusiones

El patrón respiratorio de pacientes en proceso de extubación se puede caracterizar a partir de las series temporales tiempo de espiración ( $T_E$ ), tiempo de inspiración ( $T_I$ ), duración del ciclo respiratorio ( $T_{Tot}$ ), y el índice de ventilación superficial ( $f / V_T$ ). También puede ser caracterizado a través de modelos autorregresivos AR o ARMA, con parámetros relacionados con el orden del modelo y el primer coeficiente, siendo estos los que mejor resultado han proporcionado a los clasificadores.

Se han estudiado clasificadores como la regresión logística, el análisis discriminante lineal, las máquinas de soporte vectorial, y el árbol de decisión para determinar los mejores parámetros que clasifican los diferentes grupos de pacientes. Los mejores resultados para los tres primeros métodos se obtuvieron considerando los parámetros media de  $T_E$ , primer coeficiente del modelo AR de  $T_E$  y de  $f / V_T$ . Se obtuvo una exactitud del 88%, con sensibilidades entre 93% y 96%, pero con valores de especificidad bajos.

El mejor resultado se obtuvo con el clasificador de árbol de decisión, con una exactitud del 93%, 98% de sensibilidad y 82% de especificidad, para la media y el orden del modelo de  $T_E$ ,  $T_{Tot}$ , y  $f / V_T$ .

#### Agradecimientos

Este trabajo ha sido financiado parcialmente por la ayuda TEC2010-21703-C03-01 del Ministerio de Economía y Competitividad del Gobierno de España. Los autores agradecen al equipo de los Drs. S. Benito del Hospital de la Santa Creu i Sant Pau, y A. Ballester del Hospital Universitario de Getafe su colaboración en la adquisición de las señales.

#### Referencias

- [1] J.M. Boles, J. Bion, A. Connors, M. Herridge, B. Marsh, C. Melote, R. Pearl, H. Silverman, M. Stanchina, A. Vieillard-Baron, and T. Welte. "TASKFORCE—weaning from mechanical ventilation". *Eur. Respir. J.* 29, pp. 1033–1056, 2007.
- [2] M.J. Tobin, "Advances in mechanical ventilation", *N. Engl. J. Med.*, Vol. 344, N. 26, pp. 1986-1996, 2001.
- [3] Jiin-Chyr Hsu, Yung-Fu Chen, Hsuan-Hung Lin, Chi-Hsiang Li and Xiaoyi Jiang, "Construction of Prediction Module for Successful Ventilator Weaning", *New Trends in Applied Artificial Intelligence*, pp. 766-775, 2007.
- [4] J. Sellares, I. Acerbi, H. Loureiro, R. L. Dellaca, M. Ferrer, A. Torres, D. Navajas, and R. Farre. "Respiratory impedance during weaning from mechanical ventilation in a mixed population of critically ill patients". *Br. J. Anaesth.* 103, pp. 828–832, 2009.
- [5] C. Seymour, A. Martinez, J. Christie, and B. Fuchs, "The outcome of extubation failure in a community hospital intensive care unit: a cohort study," *Critical Care* 8, R322–R327, 2004.
- [6] Conti G., L. Montini, M.A. Pennisi. "A prospective blinded evaluation of indices proposed to predict weaning from mechanical ventilation". *Intensive care Med.*, Vol. 30, No. 5, pp. 830-836, 2004.
- [7] H.R. Hemant, J. Chacko, M.K. Singh, "Weaning from mechanical ventilation- current evidence". *Indian Journal of Anaesth.* Vol. 50, No. 6, pp. 435-438, 2006.
- [8] J. Chaparro, B. Giraldo, P. Caminal, S. Benito, "Analysis of the Respiratory Pattern Variability of Patients in Weaning Process using Autoregressive Modeling Techniques". *Engineering in Medicine and Biology Society, 33<sup>rd</sup> Annual International Conference of IEEE*, pp. 5690–5693, 2011.
- [9] R.M. Oxford, L.G. Daniel, "Basic Cross-Validation: Using the "Holdout" Method To Assess the Generalizability of Results". The entity from which ERIC acquires the content, including journal, organization, and conference names, or by means of online submission from the author. *Research in the Schools*, v8 n1 pp.83-89 Spr 2001.
- [10] R. Jhonson, D. Wicher, "Applied Multivariate Statistical Analysis". 5<sup>th</sup> Edition, Prentice Hall Hispanoamericana, 2006.
- [11] C. Huberty, "Applied Discriminant Analysis, Wiley Series in Probability and Mathematical Statistics", Editorial Jhon Wiley & Sons Inc., 1994.
- [12] I. Steinwart, A. Chrismann, "Super Vector Machine, Information Science and Statistics", Ed. Springer. 2008.
- [13] A. Garde, R. Schroeder, A. Voss, P. Caminal, S. Benito and B.F. Giraldo, "Patients on weaning trials classified with support vector machines", *Physiol. Meas.* 31, pp. 979–993, 2010.
- [14] C. Burges, "A tutorial on support vector machines for pattern recognition", *Journal of Data Mining and Knowledge Discovery*, vol 1, no2. 1998.
- [15] C. Kingsford and S.L. Salzberg, "What are decision trees?". *Nat Biotechnol*, Vol. 26, No. 9, pp. 1011–1013, 2008
- [16] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases". *American Association for Artificial Intelligence*, pp. 0738-4602, 1996.
- [17] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation," *Advances in Artificial Intelligence* pp. 1015–1021, 2006.